



计算机科学

COMPUTER SCIENCE

路网拓扑感知的轨迹表示学习方法

陈嘉俊, 陈伟, 赵雷

引用本文

陈嘉俊, 陈伟, 赵雷. 路网拓扑感知的轨迹表示学习方法[J]. 计算机科学, 2023, 50(11): 114-121.

CHEN Jiajun, CHEN Wei, ZHAO Lei. [Road Network Topology-aware Trajectory Representation Learning](#) [J]. Computer Science, 2023, 50(11): 114-121.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多粒度特征融合的新型图卷积网络用于方面级情感分析](#)

Novel Graph Convolutional Network Based on Multi-granularity Feature Fusion for Aspect-based Sentiment Analysis

计算机科学, 2023, 50(10): 80-87. <https://doi.org/10.11896/jsjcx.230600036>

[融合音字特征转换的非自回归Transformer中文语音识别](#)

Non-autoregressive Transformer Chinese Speech Recognition Incorporating Pronunciation-Character Representation Conversion

计算机科学, 2023, 50(8): 111-117. <https://doi.org/10.11896/jsjcx.220600144>

[基于改进粒子群算法的云数据中心能耗优化任务调度策略](#)

Task Scheduling Strategy for Energy Consumption Optimization of Cloud Data Center Based on Improved Particle Swarm Algorithm

计算机科学, 2023, 50(7): 246-253. <https://doi.org/10.11896/jsjcx.220900176>

[语义风格一致的任意图像风格迁移](#)

Arbitrary Image Style Transfer with Consistent Semantic Style

计算机科学, 2023, 50(7): 129-136. <https://doi.org/10.11896/jsjcx.220700008>

[基于改进Yolov4-tiny的轻量型目标检测算法](#)

Lightweight Target Detection Algorithm Based on Improved Yolov4-tiny

计算机科学, 2023, 50(6A): 220700006-7. <https://doi.org/10.11896/jsjcx.220700006>

路网拓扑感知的轨迹表示学习方法

陈嘉俊 陈伟 赵雷

苏州大学计算机科学与技术学院 江苏 苏州 215006

(20204227012@stu.suda.edu.cn)

摘要 现有路网场景下的轨迹表示学习 (Trajectory Representation Learning, TRL) 方法可分为两类,即基于循环神经网络 (RNN) 和长短期记忆 (LSTM) 的序列化模型以及基于自注意力机制的学习模型。尽管已有研究做出了重大贡献,但它们仍然存在以下问题:(1) 现有的路网表示学习方法忽略了相邻路段之间的转移概率,不能充分捕获路网的拓扑结构信息;(2) 基于自注意力机制的学习模型在短轨迹和中长轨迹上的表现优于序列化模型,但在长轨迹的表示学习上性能较差,未能很好刻画轨迹的长期语义特征。基于此,文中提出了一个新的轨迹表示学习模型 TRMS。该模型采用概率感知游走来优化传统 DeepWalk 算法,以深入挖掘路网的拓扑结构,然后将自注意力机制和 Masked Seq2Seq 学习框架相结合来捕获轨迹的长期语义特征。最后,基于真实轨迹数据进行实验,结果表明,TRMS 在短、中、长轨迹的嵌入表示上,性能都优于最好的基线方法。

关键词: 路网; 拓扑结构; 轨迹表示学习; 序列化模型; 自注意力机制

中图分类号 TP311

Road Network Topology-aware Trajectory Representation Learning

CHEN Jiajun, CHEN Wei and ZHAO Lei

School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu 215006, China

Abstract The approaches developed for task trajectory representation learning (TRL) on road networks can be divided into the following two categories, i. e., recurrent neural network (RNN) and long short-term memory (LSTM) based sequence models, and the self-attention mechanism based learning models. Despite the significant contributions of these studies, they still suffer from the following problems. (1) The methods designed for road network representation learning in existing work ignore the transition probability between connected road segments and cannot fully capture the topological structure of the given road network. (2) The self-attention mechanism based learning models perform better than sequence models on short and medium trajectories but underperform on long trajectories, as they fail to character the long-term semantic features of trajectories well. Motivated by these findings, this paper proposes a new trajectory representation learning model, namely trajectory representation learning on road networks via masked sequence to sequence network (TRMS). Specifically, the model extends the traditional algorithm DeepWalk with a probability-aware walk to fully capture the topological structure of road networks, and then utilizes the Masked Seq2Seq learning framework and self-attention mechanism in a unified manner to capture the long-term semantic features of trajectories. Finally, experiments on the real-world datasets demonstrate that TRMS outperforms the state-of-the-art methods in embedding short, medium, and long trajectories.

Keywords Road-network, Topological structure, Trajectory representation learning, Sequence model, Self-attention mechanism

1 引言

随着移动终端在日常生活中的普及,大规模轨迹数据正以前所未有的速度生成^[1],而且许多轨迹都诞生于路网。值得注意的是,这些轨迹不仅包含丰富的路网拓扑结构信息,还包含移动对象的序列化行为信息。基于这些数据,可进一步探索下一个位置预测^[2]、终点预测^[3]、轨迹相似性计算^[4]和

异常点检测^[5]等下游任务。

作为上述研究的重要组成部分,近年来轨迹表示学习 (TRL) 受到了越来越多研究者的关注。直观地讲,TRL 的目的是将原始轨迹数据 (即采样点序列) 转换为适合后续任务的通用格式 (例如向量)。传统的 TRL 方法大多依赖于人工设计的特征,主要依靠领域专家的先验知识和经验,且仅对有限数量的特征进行探索,如最大加速度速度^[6]、变化率^[7]、停止

到稿日期:2022-10-09 返修日期:2023-02-13

基金项目:国家自然科学基金(61902270);江苏省高等学校基础科学(自然科学)研究重大项目(19KJA610002)

This work was supported by the National Natural Science Foundation of China(61902270) and Major Program of the Natural Science Foundation of Jiangsu Higher Education Institutions of China(19KJA610002).

通信作者:陈伟(robertchen@suda.edu.cn)

率^[8]、转向率^[8]等。近年来,得益于深度学习模型强大的信息表征能力,越来越多的研究者开始将神经网络应用到 TRL 任务。其中,一些研究者用 mask 向量表示路网拓扑结构对轨迹的约束,然后基于序列化模型对轨迹进行表征,如 CS-RNN^[9], Trembr^[10]等;另一些研究者则设计了基于自注意力机制的表示学习模型,如 Toast^[11]。尽管上述研究作出了重要贡献,但由于以下两方面的原因,他们的工作针对路网场景下的轨迹表示学习仍存在一定的局限性。

首先,对于路网表示学习,现有模型如 HRNR^[12], RFNs^[13], IRN2Vec^[14]等,通常是在路网上基于随机游走获取训练数据,忽略了路段与其相邻路段间潜在的转移概率。然而,在现实生活中,这个概率对预测移动对象的运动倾向是至关重要的。通常,不同路段间的转移概率是不同的,概率越大,说明路段间潜在的内部关联越强。例如,在市中心的十字路口,车辆转向主干道的概率通常比转向小路的概率大。遗憾的是,现有研究忽略了路网表示学习中这一重要信息,未能充分捕获路网的拓扑结构信息。虽然文献[12]试图通过考虑路段类型来提高路网表示学习的准确性,如果不考虑路段间的转移概率,模型性能的提升是有限的。

其次,现有 TRL 方法难以有效表征不同长度轨迹的语义特征,特别是对现实生活中广泛存在的长轨迹(如居住在郊区但在市中心工作的人每天都会产生大量长轨迹)的特征表征,模型的准确率很低。与短轨迹相比,蕴含更多时空信息的长轨迹更能反映移动对象的行为模式,而目前最优的基于自注意力机制的模型 Toast^[11]在中、短轨迹上表现突出,但在长轨迹上表现不佳,未能很好地刻画轨迹的长期语义特征。其原因如下:Toast 首先以一定的概率(如 15%)对轨迹对应路径进行随机掩码,然后利用自注意力机制挖掘路段之间潜在的内部联系。然而,在已知相邻路段的情况下,模型可能只依靠路网拓扑结构的先验知识就能推断出掩码路段,致使模型没有充分挖掘轨迹的序列化行为信息,难以有效表征长期语义特征。因此,为了提高模型刻画轨迹长期语义特征的有效性,我们需要同时考虑路网的拓扑结构和轨迹的序列化特征。

针对现有研究的不足,本文提出了一种新的针对路网场景下不同长度轨迹的表示学习模型 TRMS(Trajectory Representation Learning on Road Networks via Masked Sequence to Sequence Network)。具体来讲,本文设计了模块 TDW(Traffic-enhanced DeepWalk),用于充分挖掘路网的拓扑结构。不同于以往的研究,该模块基于概率感知游走对传统 DeepWalk^[15]算法进行了优化。首先基于大量历史轨迹数据,根据相邻路段出现的频率计算路段间的转移概率,然后在路网上基于概率感知游走来捕获拓扑信息。此外,为了有效表征轨迹的长期语义特征,本文基于文献[16]设计了模块 MSN(Masked Sequence to Sequence Network),该模块从每条轨迹对应路径的随机位置开始,在编码阶段连续掩码长度为 k 的片段,接着训练 MSN 对掩码片段进行预测,这样模型可以充分挖掘轨迹的序列化特征。本文的主要贡献如下:

(1)为有效表征路网场景下不同长度的轨迹,设计了 TRMS。该模型同时考虑了路网的拓扑结构和轨迹的长期语义特征。

(2)基于概率感知游走,设计了 TDW 以充分挖掘路网的拓扑结构;结合 Masked Seq2Seq 学习框架和自注意力机制开发了 MSN,以更好地表征轨迹的长期语义特征。

(3)为了验证 TRMS 的有效性,基于真实轨迹数据进行了大量实验,结果表明,在路网场景下不同长度轨迹的特征表征上,TRMS 优于现有最好的基线方法。

2 相关工作

2.1 路网表示学习

路网表示学习的目的是挖掘路段间潜在的内部关联。现有研究通常是将路网视为一张有向图,其中边代表路段,顶点代表路口,通过学习节点或边缘的有效特征表示来表征路网。一些方法基于空间域提取图中的节点特征,如 DeepWalk 和 Node2Vec^[17],它们首先在路网上随机游走获得路段序列,然后基于 Skip-gram^[18]学习路段表示。还有一些方法试图基于谱域来对图中的节点进行特征表征,如基于 GCN 的 HRNR, RFNs 和 IRN2Vec,它们通过聚合当前节点的邻居节点特征来获得其嵌入表示。上述研究虽取得了不错的成果,但忽略了相邻路段间可用于更充分挖掘路网拓扑结构的转移概率。

2.2 轨迹表示学习

TRL 的目的是将轨迹转化为低维特征向量,以从原始冗余数据中提取出有价值的信息。现有研究通常采用自然语言处理中的方法来挖掘轨迹的序列化特征。常见序列化建模方法有基于 RNN, LSTM 和 GRU 的序列化模型以及基于自注意力机制的模型,如 BERT^[19]和 Transformer^[20]。具体来说,CSRNN 采用 mask 向量表示路网拓扑结构对轨迹的约束,通过将 mask 向量输入模型 softmax 层来对轨迹进行特征表征。Trembr 首先采用 Skip-gram 来获得路段的嵌入表示,然后基于 LSTM 学习轨迹的特征表示。模型 t2vec^[4]使用网格对地图空间进行划分,将轨迹转换为网格单元序列,然后基于 GRU 对网格序列建模,但其性能受限于网格大小的划分。此外,文献[11]结合 BERT 开发了基于自注意力机制的模型 Toast,以对路网场景下的轨迹进行建模。但正如第 1 章所讨论的,以上工作不能很好地表征不同长度的轨迹,难以有效捕获轨迹的长期语义特征信息。

3 问题描述

本章首先介绍本文中使用的几个核心定义,然后阐述本文的研究问题。相关符号含义如表 1 所列。

表 1 相关符号含义

Table 1 Meaning of symbols

Symbol	Meaning
V	顶点集合
E	路段集合
T	一条轨迹
D_T	轨迹数据集
R	一个路段
R	一条路径
D_R	路径数据集
M	转移概率矩阵
G	路网
s	一个概率感知游走序列
S	概率感知游走序列集

定义 1 (路网) 路网被定义为一张有向图 $G=(V, E, M)$, 其中 V 是顶点集合(如路口), E 是边集合(一条边代表一个路段), $M \in \mathbb{R}^{|E| \times |E|}$ 是基于大量历史轨迹计算出的转移概率矩阵。

定义 2 (轨迹) 一条轨迹是由移动对象在路网上生成的采样点序列, 定义为 $T=(p_1, p_2, \dots, p_T)$, 其中每个采样点 p_i 对应一个经纬度坐标。

定义 3 (路径) 一条路径 $R=(r_1, r_2, \dots, r_R)$ 代表一条相邻路段序列, 其中 $r_i \in E$ 表示给定路网上的某一路段。

为消除轨迹采样点采样率低且不均匀以及噪声干扰点的影响^[21-22], 本文通过文献[23]提出的 map-matching 算法将每条轨迹 T 映射到路网 G 上, 得到相应路径 R , 并将路径集记为 D_R 。如图 1(a)所示, 轨迹 T_1 转换为图 1(b)中的路径 $R_1=(r_1, r_2, r_3, r_4)$ 。

问题定义如下: 给定一张路网 $G=(V, E, M)$ 和一个包含不同长度轨迹的数据集 D_T , 本文研究的最终目标是训练一个自监督模型来学习 D_T 中每条轨迹 T 的嵌入表示, 即一个 d_{model} 维向量。该特征向量既包含路网的拓扑结构信息, 也

包含轨迹的语义特征信息, 进而使得模型在这个向量空间上能有效解决常见的轨迹下游任务。

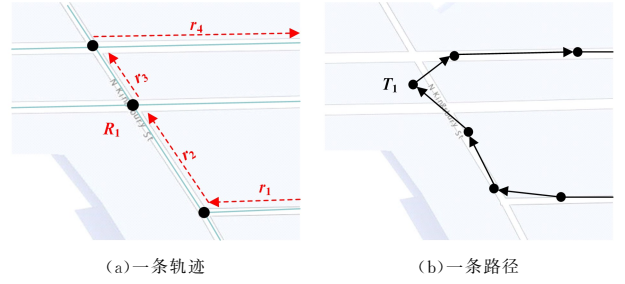


图 1 路网场景下的一条轨迹及其对应路径

Fig. 1 Example of a trajectory and corresponding route on road networks

4 模型 TRMS 的结构

TRMS 的框架如图 2 所示, 本章首先介绍了 TDW, 然后阐述了 MSN 的各个组成部分, 最后提出了本文针对路网场景下的轨迹表示学习所设计的预训练任务。

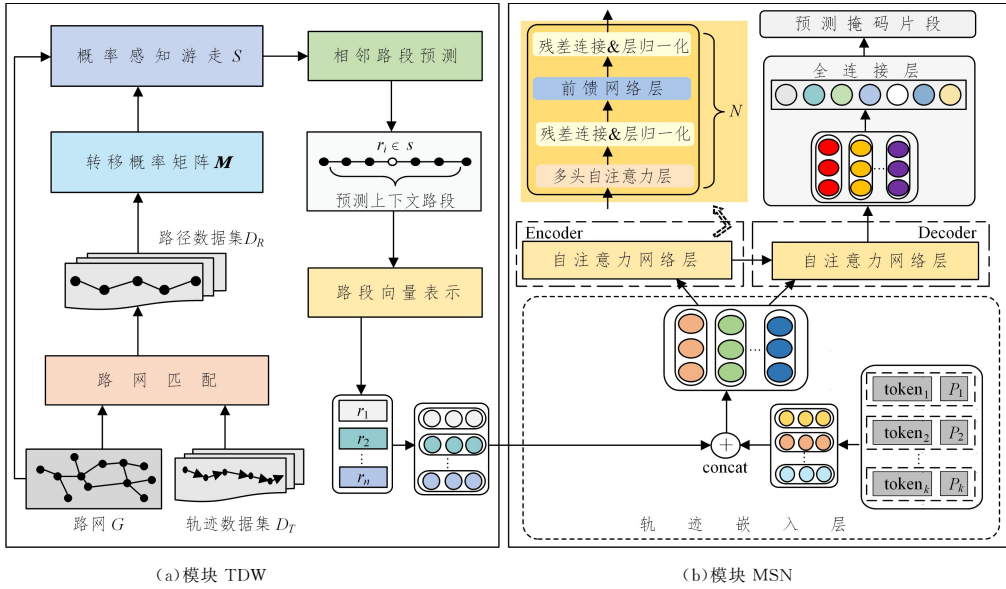


图 2 TRMS 的模型框架

Fig. 2 Overview of the proposed model TRMS

4.1 模块 TDW

4.1.1 路网匹配

本文首先对原始轨迹数据进行预处理, 对噪声采样点采取删除和拟合两种方式来提升轨迹数据的质量; 然后给定一组历史轨迹和路网 G , 为了减小数据稀疏性对路网匹配的影响, 我们采用文献[23]提出的基于隐马尔可夫模型的方法, 将每条轨迹 T 映射到路网上, 得到对应路径 R 。

4.1.2 路段表示学习

路网匹配完成后, 得到路径集 D_R , 然后构造转移概率矩阵 M , 其中每个元素 $M[i, j]$ 表示路径 $r_i \rightarrow r_j$ 的概率, 计算式如式(1)所示:

$$M[i, j] = \frac{N[i, j]}{\sum_{i=1}^{|V|} N[i, j]} \quad (1)$$

其中, $N[i, j]$ 表示路径 $r_i \rightarrow r_j$ 在历史轨迹中出现的次数, $N[i,$

$j] > 0$, 当且仅当路段 r_i 和路段 r_j 相邻, 且在某条轨迹上至少一起出现过一次。

在构造转移概率矩阵 M 之后, 接着介绍概率感知游走的过程: 首先在路网中随机选择一个路段 r_i 作为起点, 然后基于 $M[i, :]$ 中的概率选择下一跳路段, 这样做是为了充分利用路段间潜在的内部关联。重复该选择过程, 可以得到一个概率感知游走序列 $s_i = \{r_i, \dots, r_k\}$, 将所有游走序列的长度都设置为一个固定值。最后, 基于 Skip-gram 方法对路段进行嵌入表示。TDW 通过最小化以下损失函数来进行优化:

$$L_{\text{TDW}} = - \sum_{r_i \in s_i} \sum_{r_j \in C(r_i)} \log P(r_j | r_i) = - \sum_{r_i \in s_i} \sum_{r_j \in C(r_i)} \log P \left(\frac{\exp(\mathbf{r}_i^T \mathbf{r}_j)}{\sum_{r_j' \in V} \exp(\mathbf{r}_i^T \mathbf{r}_j')} \right) \quad (2)$$

其中, s_i 是一个概率感知游走序列, $C(r_i)$ 代表路段 r_i 在 s_i 中的上下文路段集合, \mathbf{r}_i 和 \mathbf{r}_j 分别为路段 r_i 和 r_j 的嵌入表示。通过

优化该损失函数,TDW 可以深入挖掘相邻路段间潜在的内部联系,以有效表征路网的拓扑结构。

4.2 模块 MSN

4.2.1 模块架构

得益于堆叠双向自注意力网络在序列化信息建模方面的强大功能,以及在长文本序列生成任务上 Masked Seq2Seq 优于 BERT,GPT^[24] 和 Transformer 的出色表现,受此启发,本文设计了 MSN 以有效表征轨迹的长期语义特征。此外,与文献[16]提出的方法 MASS 相比,MSN 主要优势在于它是基于双向多层自注意力网络设计的,可以根据未掩码片段的上下文信息,来提取出轨迹的语义特征,从而精准预测出掩码片段。接下来,将以自底向上的方式介绍 MSN,最后详细描述本文为 TRL 所设计的预训练任务。

(1) 轨迹嵌入层。针对路径中的每个路段 r_i , 首先初始化其嵌入表示 $token_i$, 其次由式(3)得到其相应的位置嵌入表示 p_i , 最后将 TDW 学习到的 r_i 与 $token_i$ 和 p_i 的和拼接, 得到轨迹嵌入层的输出 e_i , 其计算式如下:

$$\begin{aligned} e_i &= (r_i \oplus (p_i + token_i))W_0 + b_0 \\ p_i(2j) &= \sin\left(\frac{i}{10000^{2j/d_{\text{model}}}}\right) \\ p_i(2j+1) &= \cos\left(\frac{i}{10000^{(2j+1)/d_{\text{model}}}}\right) \end{aligned} \quad (3)$$

其中, d_{model} 为轨迹嵌入维数, W_0 和 b_0 为轨迹嵌入层参数。接着, 将 e_i 作为双向自注意力网络层的输入, 过程如图 2(b) 所示。

(2) 自注意力网络层。近年来, 自注意力机制在图像处理、语音识别和机器翻译等领域得到了广泛的应用, 它减少了序列对外部信息的依赖, 专注于挖掘序列内部数据间的相互作用, 并且在处理长序列数据时具有并行计算的优势。形式上, 它被定义为:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

其中, Q, K 和 V 分别是维度为 d_q, d_k 和 d_v 的查询向量、键向量和值向量。

给定一条长度为 n 的路径 $R = (r_1, \dots, r_n)$, 首先经过轨迹嵌入层, 得到其嵌入表示 $R = [e_1, \dots, e_n] \in \mathbb{R}^{n \times d_{\text{model}}}$, 然后经自注意力网络层得到其输出表示 $H = [h_1, \dots, h_n] \in \mathbb{R}^{n \times d_{\text{model}}}$ 。其中, 多头自注意力机制是将路径中的每个路段嵌入表示映射到 h 个子向量空间中, 然后从不同方面深入挖掘路段间潜在的关联, 每个映射空间表示成一个“head”。通过整合所有子空间的特征信息, 得到以下输出:

$$\begin{aligned} head_i &= Attention(RW_i^Q, RW_i^K, RW_i^V) \\ H &= MultiHead(R) = [head_1, \dots, head_h]W^O \end{aligned} \quad (5)$$

其中, $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d_{\text{model}} \times d_{\text{model}}/h}$, $W^O \in \mathbb{R}^{d_{\text{model}} \times d_{\text{model}}}$ 是自注意力网络层参数。

(3) 前馈网络层。将自注意力网络层的输出 H 输入到由两个全连接层构成的前馈网络层中, 该前馈网络层的输出 Y 由式(6)计算得到:

$$PFN(Y) = \max(0, HW_1 + b_1)W_2 + b_2 \quad (6)$$

其中, W_1, W_2, b_1, b_2 分别为两个全连接层参数。

(4) 残差连接与层归一化。为避免多层神经网络叠加带来的梯度消失等问题, 本文在自注意力网络层与前馈网络层之间应用残差连接, 其表达式为:

$$\begin{aligned} Y' &= LayerNorm(R + MultiHead(R)) \\ R' &= LayerNorm(Y' + PFN(Y')) \end{aligned} \quad (7)$$

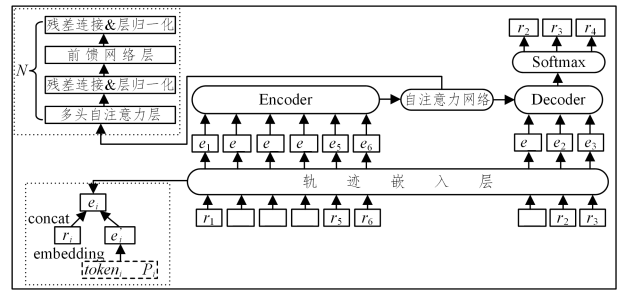
其中, $LayerNorm$ 表示层归一化; Y' 表示当前网络层的输出, 同时其也作为下一网络层的输入。

4.2.2 模型预训练

对于给定的路径 $R \in D_R$, 将 $R^{i:j}$ 作为 R 的 mask 形式, 其中 R 从位置 i 到 j ($0 < i < j < n$) 的片段被掩码, 对应长度记为 $k = j - i + 1$, n 为路径 R 的长度, 将 $R^{i:j}$ 用来表示 R 被掩码前从位置 i 到 j 的原始路径片段。本文用一个特殊符号 [Mask] 表示每个掩码路段。接下来, 将 $R^{i:j}$ 作为 Encoder 的输入, 将 $R^{i:j}$ 作为 Decoder 的输入, 训练 MSN 预测出掩码片段。最后, 使用对数似然作为目标函数进行优化。

$$\begin{aligned} L(\theta; D_R) &= \frac{1}{|D_R|} \sum_{R \in D_R} \log P(R^{i:j} | R^{i:j}; \theta) \\ &= \frac{1}{|D_R|} \sum_{R \in D_R} \log \prod_{t=i}^j P(R_t^{i:j} | R_{<t}^{i:j}, R^{i:j}; \theta) \end{aligned} \quad (8)$$

如图 3 所示, Encoder 的输入有 6 个路段, 其中路段 r_2, r_3, r_4 被掩码。MSN 在预训练过程中需捕获轨迹语义信息以预测出这些连续掩码片段。



注: “_”表示掩码符[Mask]。

图 3 MSN 的 encoder-decoder 框架

Fig. 3 Encoder-decoder framework of MSN

具体来讲, MSN 在编码过程中从每条路径随机位置开始, 连续掩码长度为 k 的片段, 并根据 TDW 学习到的路段表示将每条路径转换为等维嵌入矩阵, 并将其输入到自注意力网络层, 接着训练 Encoder, 从剩余未掩码片段中提取轨迹的语义特征, 同时模型采用自注意力机制进一步挖掘路径内部各路段间的相互关联, 最后将这些信息传递给 Decoder 以预测出掩码片段。这样 MSN 可以充分挖掘轨迹序列化行为信息, 更好地表征轨迹的长期语义特征。此外, 为了平衡 MSN 中 Encoder 和 Decoder 的性能, 本文将 k 设为每条路径长度的 50%。在 5.5 节中, 通过改变参数 k 的值来进行分析, 并给出相应的 TRMS 性能结果。算法 1 描述了 TRMS 的训练过程。

算法 1 TRMS 的训练过程

输入: 路网 $G = (V, E, M)$, 轨迹集 D_T , 训练次数 n , 目标和邻居路段嵌入 r_i 和 $r_j, i, j \in [1, |E|]$

输出: 每个路段 r 和路径 R 的特征表示 r 和 R

1. 将每条轨迹 $T \in D_T$ 映射到 G 上得到路径集 D_R ;
2. 根据 D_R 计算得到转移概率矩阵 M ;

```

3. for  $n \leftarrow 1$  to  $n$  do:
4.   在  $G$  上生成一个概率感知游走序列集  $S$ ;
5.   for  $s \in S$  do:
6.     反向传播最小化  $L_{TDW}$  来更新  $r_i$  和  $r_j$ ;
7.   end for;
8.   for  $R \in D_R$  do:
9.     执行 4.2.2 节所介绍的预训练任务;
10.    计算预训练任务的损失函数  $L(\theta; D_R)$ ;
11.    反向传播最小化  $L(\theta; D_R)$  来更新  $r$  和  $R$ ;
12.   end for;
13. end for;

```

此外,受机器翻译的启发,本文在每条路径开头插入一个占位符,并将其作为最终模型学习到的轨迹表示。我们将学到的轨迹表示输入到具体的下游任务模块,以验证 TRMS 能够有效表征路网场景下不同长度的轨迹。

5 实验

5.1 数据集

本文使用北京和波尔图两个城市的真实路网数据和轨迹数据。其中,路网数据采样自 OpenStreetMap,轨迹数据来源于 2012 年 1 月 10 日至 2012 年 10 月 18 日期间北京的 3025 辆出租车,以及 2013 年 1 月 7 日至 2014 年 6 月 30 日期间波尔图的 442 辆出租车。为了验证 TRMS 能够有效表征不同长度的轨迹,本文根据每条路径所涉及的路段数量,将原始轨迹数据集划分成 6 个子轨迹数据集,分别是北京的 BJ-Short, BJ-Medium, BJ-Long 和波尔图的 PT-Short, PT-Medium, PT-Long。这 6 个子轨迹数据集的统计情况如表 2 所列。

表 2 预处理后的 6 个子轨迹数据集的统计情况

Table 2 Statistics of six trajectory datasets after preprocessing

数据集	路段数	轨迹数	平均轨迹长度/km	路径长度区间
BJ-Short	7266	586627	3.35	小于 40
BJ-Medium	7254	293954	6.44	40~100
BJ-Long	7075	71229	12.25	大于 100
PT-Short	2988	723892	3.60	小于 40
PT-Medium	3006	361957	6.66	40~100
PT-Long	2949	82563	13.44	大于 100

5.2 验证任务与度量指标

为了验证 TRMS 学习到的嵌入表示能够有效刻画路网的拓扑结构和轨迹的长期语义特征,本文采用下一个位置预测、终点预测和轨迹相似度查询作为 3 个下游任务。

(1)下一个位置预测。位置预测是许多移动应用程序的基础任务,如推荐系统和交通规划。该任务的目的是预测移动对象下一个最有可能经过的位置。本文将一个路段视为一个位置,将每条路径的最后一个路段作为预测目标。TRMS 经过训练,使得各个路段的相邻路段在其下一个位置候选集中具有较高的排名。

(2)终点预测。终点预测是基于局部轨迹去预测最终目的地。与下一个位置预测任务类似,在该项工作中,我们将路径的最后 4 个路段掩码后输入模型,将最后一个路段视为终点预测目标。同样地,TRMS 经过学习,需要准确感知移动对象真实的行为模式,并预测出其最终目的地。

(3)轨迹相似度查询。本文基于嵌入表示来计算轨迹间在向量空间上的相似性得分。此外,我们还将 TRMS 与 LCSS,EDR 和 Fréchet 这 3 种常用轨迹相似性度量方法进行比较。在这项研究中,我们随机抽取了 5000 条轨迹作为查询集。为验证不同方法的有效性,我们为每条查询轨迹生成一条相似轨迹,并将其作为查询轨迹的 Ground-truth。具体来说,对于给定的一条路径,从中随机选择一个开始路段和结束路段,然后沿着偏离原路线方向的另一条路线在这两个路段之间绕行,平均绕行长度为整个路径的 6.3%。接着又随机抽取了 2.5 万条轨迹,将其与前面生成的 5000 条 Ground-truth 一起作为总查询数据库。

(4)度量指标。对于下一个位置预测和终点预测任务,TRMS 尝试为每个测试集轨迹预测其概率最大的前 K 个候选路段。本文使用预测精度 $ACC@1$ 和 $ACC@5$ 作为度量指标。 $ACC@K$ 被定义为:

$$ACC@K = \frac{\# \text{correctly predicted trajectories } @K}{\# \text{trajectories}} \quad (9)$$

对于轨迹相似度查询任务,本文使用 Mean Rank (MR) 和 $HR@10$ 作为度量指标。MR 指在查询过程中,每条轨迹的 Ground-truth 在模型预测结果中的平均排名; $HR@10$ 指模型预测出的前 10 条最相似轨迹中,包含查询轨迹的 Ground-truth 的命中率。

(5)参数设置。对于所有基线方法,本文将参数调整为各自最优的情况。本文将路段和轨迹的嵌入维度都设置为 256。在 TDW 中,以每个路段为起点将概率感知游走的次数、游走序列长度和 Skip-gram 窗口大小分别设置为 15,40 和 5。自注意力网络层堆叠 3 层, K, Q, V 维度均设为 64,head 设为 4。模型参数采用 SGD 优化器,学习率设为 0.001,动量设为 0.99。

5.3 基线方法

需要注意的是,本文研究的目的是在相同条件下,用其他方法来验证 TRMS 的有效性,因此没有采用针对特定任务所设计的复杂模型或使用其他额外数据源。本文采用以下方法作为对比方法。

(1)CSSRNN。该模型基于序列化模型 RNN,将路网的拓扑约束以 mask 向量的形式输入 softmax 层,以完成对轨迹的嵌入表示。

(2)t2vec。该模型针对轨迹相似性计算任务,设计了一个 Seq2Seq 学习框架用于对轨迹建模,其 Encoder 和 Decoder 端都采用双向 LSTM 网络。

(3)Trembr。该模型首先基于词嵌入的方法获得路段表示,然后使用基于 LSTM 的 Encoder-Decoder 框架来学习轨迹表示。

(4)HRNR-MSN。该模型将文献[12]中的路网表示学习模型 HRNR 与本文设计的 MSN 相结合,以验证 TDW 的有效性。

(5)Toast。这是目前研究路网场景下轨迹表示学习的最优方法,它结合 BERT 开发了一种基于自注意力机制的方法,用于刻画轨迹的语义特征和路网的拓扑结构。

此外,本文在消融实验部分将传统 Deepwalk 方法替换为

TDW,以研究 TDW 对模型性能的贡献。

5.4 实验结果与分析

首先,对于轨迹相似度查询任务,本文使用几种常用测量方法来比较和评估 TRMS 的性能。由于几个数据集的实验结果相似,因此本文只给出模型在中长轨迹数据集上的结果。

实验结果如表 3 所列。我们可以看出,TRMS 在两个评估指标上都优于现有基线方法,这表明 TRMS 学习到的嵌入表示能够有效表征路网场景下不同长度的轨迹特征,让相似轨迹在同一向量空间上距离相近。

表 3 轨迹相似度查询的实验结果

Table 3 Results of trajectory similarity search

	BJ-Medium		PT-Medium	
	MR	HR@10	MR	HR@10
LCSS	71.49	0.453	64.41	0.482
EDR	461.63	0.164	455.20	0.196
Fréchet	25.27	0.802	22.67	0.827
t2vec	49.38	0.764	47.84	0.782
Toast	12.10	0.847	11.47	0.863
TRMS	8.56	0.875	7.28	0.901

接着,对于下一个位置预测和终点预测任务,实验结果如

表 4 下一个位置预测的性能比较

Table 4 Performance comparison of next location prediction

Task	BJ-Short		BJ-Medium		BJ-Long		PT-Short		PT-Medium		PT-Long	
	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5
CSRNN	0.778	0.914	0.793	0.931	0.563	0.714	0.821	0.954	0.838	0.954	0.672	0.860
t2vec	0.786	0.933	0.814	0.941	0.587	0.728	0.830	0.965	0.852	0.971	0.694	0.874
Trembr	0.804	0.946	0.825	0.949	0.562	0.701	0.842	0.969	0.857	0.976	0.688	0.865
HRNR-MSN	0.812	0.952	0.833	0.854	0.614	0.772	0.85	0.974	0.871	0.983	0.731	0.882
Toast	0.823	0.949	0.841	0.943	0.541	0.694	0.853	0.981	0.879	0.975	0.617	0.813
TRMS	0.847	0.981	0.867	0.978	0.653	0.807	0.878	0.992	0.907	0.995	0.763	0.917

表 5 终点预测的性能比较

Table 5 Performance comparison of destination prediction

Task	BJ-Short		BJ-Medium		BJ-Long		PT-Short		PT-Medium		PT-Long	
	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5	ACC@1	ACC@5
CSRNN	0.564	0.756	0.611	0.823	0.389	0.571	0.659	0.877	0.695	0.857	0.544	0.748
t2vec	0.578	0.783	0.634	0.845	0.384	0.566	0.683	0.890	0.711	0.874	0.567	0.761
Trembr	0.604	0.813	0.647	0.853	0.397	0.575	0.699	0.904	0.736	0.892	0.559	0.753
HRNR-MSN	0.624	0.842	0.651	0.857	0.425	0.642	0.708	0.909	0.752	0.913	0.573	0.772
Toast	0.633	0.849	0.662	0.871	0.376	0.562	0.715	0.913	0.763	0.925	0.452	0.654
TRMS	0.651	0.879	0.687	0.902	0.463	0.674	0.736	0.947	0.794	0.960	0.614	0.805

尽管以上方法在长轨迹嵌入表示上的 ACC@1 和 ACC@5 都不是很高,但 TRMS 在 BJ-Long 和 PT-Long 上的性能仍然有最大的提升。

5.5 模型分析

5.5.1 消融实验

TRMS 由两个模块组成,即 TDW 和 MSN。为了说明 TDW 和 MSN 对模型性能的影响,本节将它们分别独立去除并基于 ACC@1 指标记录相关实验结果。本文准备了 4 种 TRMS 的变体进行比较,分别是:(1)TRMS-LM,该模型将 MSN 替换为基于 LSTM 的 Seq2Seq 网络;(2)TRMS-BT,该模型将 TDW 与自注意力网络 BERT 相结合;(3)TRMS-DM,该模型用传统 Deepwalk 方法代替 TDW;(4)TRMS-MS,该模型仅使用 MSN 对轨迹进行建模。

所有消融实验的结果如图 4—图 6 所示,从中可以观察到:

表 4 和表 5 所列。由表 4、表 5 可观察到:

(1)基于自注意力机制的模型 Toast 在中、短轨迹建模方面优于序列化模型 CSSRNN,t2vec 和 Trembr,但在长轨迹建模方面表现不佳。一方面,这表明自注意力机制能够有效挖掘路段间潜在的内部联系;另一方面,这也反映了其不能很好地表征轨迹的长期语义特征。

(2)基于 ACC@1 指标,TRMS 在短、中、长轨迹嵌入上分别优于 HRNR-MSN 2.95%,3.7%和 3.75%。这说明相比 HRNR,TDW 能更有效地挖掘路网的拓扑结构信息。

(3)基于 ACC@1 指标,TRMS 在嵌入短、中、长轨迹方面分别优于 Toast 2.2%,2.75%和 12.68%。这表明 TRMS 能够有效表征轨迹的长期语义特征,原因是 MSN 在编码过程中从每条路径随机位置开始,连续掩码长度为 k 的片段,接着训练 Encoder 从剩余未掩码片段中提取轨迹语义特征,同时模型采用自注意力机制进一步挖掘路径内部各路段间的相互关联,最后将这些信息传给 Decoder 以对掩码片段进行预测,这样 MSN 可以充分捕获轨迹的序列化行为信息。由于 TRMS 同时考虑了路网的拓扑结构和轨迹的语义信息,因此可以提高轨迹预测的精度。

(1)TRMS-MS 在预测任务上的准确率低于 TRMS,这说明路网场景下的轨迹表示学习确实需要考虑路网拓扑结构的约束,而 TDW 能有效捕获这一拓扑信息,进而帮助 TRMS 更好地进行轨迹的特征表征。

(2)TRMS 的性能优于 TRMS-DM,这说明在路网上基于概率感知游走而不是随机游走,该方式能更充分地挖掘路段之间潜在的内部联系,进而提高路段表示学习的有效性。

(3)与 TRMS-LM 相比,TRMS 在短、中、长轨迹嵌入方面的表现更好,这表明基于自注意力机制的方法优于基于序列化模型的方法。具体来讲,TRMS 相比 TRMS-LM 在 ACC@1 上平均高出 5.08%,在 ACC@5 上平均高出 6.17%。

(4)与 TRMS-BT 相比,TRMS 在不同长度的轨迹建模上都具有更高的准确率,尤其是在长轨迹建模方面,TRMS 在 ACC@1 和 ACC@5 指标上分别优于 TRMS-BT 11.2%和

13.75%，这说明 MSN 能有效刻画轨迹的长期语义特征。正如第 1 章所述，以一定概率对轨迹对应路径进行随机掩码，会造成模型仅凭路网拓扑结构的先验知识，根据相邻路段就推断出掩码路段，致使模型没有学习到隐藏在轨迹数据中的语义特征。

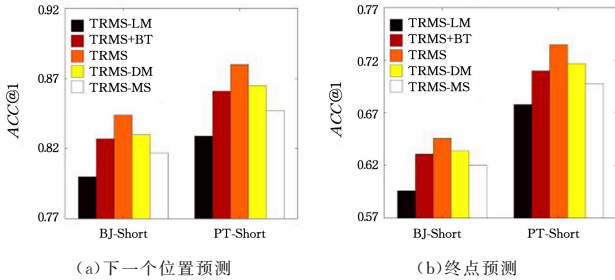


图 4 TRMS 在短轨迹集上的消融研究

Fig. 4 Ablation study of TRMS on short trajectories

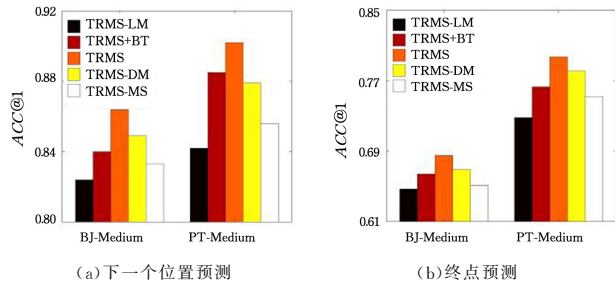


图 5 TRMS 在中长轨迹集上的消融研究

Fig. 5 Ablation study of TRMS on medium trajectories

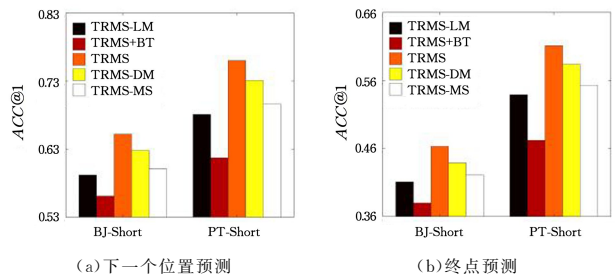


图 6 TRMS 在长轨迹集上的消融研究

Fig. 6 Ablation study of TRMS on long trajectories

综上所述，TDW 和 MSN 都对 TRMS 的性能做出了重要贡献，通过结合这两部分，所提模型可以有效表征路网场景下不同长度的轨迹。

5.5.2 参数分析

为了获得 TRMS 的最佳性能，参数调优至关重要。由于篇幅所限，本文只报告轨迹嵌入维度 d_{model} 对下一个位置预测任务的影响，以及编码过程中掩码长度 k 对终点预测任务的影响。

(1) 轨迹嵌入维度 d_{model} 。如图 7 所示，将北京数据集和波尔图数据集轨迹的嵌入维度设置为 256 时，模型性能达到最佳。通常来说，当轨迹较短时，嵌入维度需要设置得小一些，当轨迹较长时，嵌入维度需要设置得较大些。嵌入维度过小，模型不能充分挖掘轨迹的语义特征和路网的拓扑结构，而嵌入维度过大，可能会引入噪声信息，导致模型性能呈下降趋势。

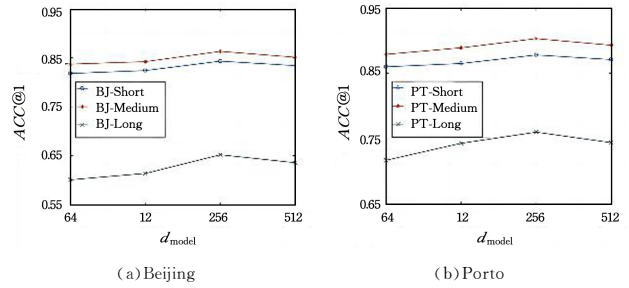


图 7 参数 d_{model} 的分析

Fig. 7 Analysis of parameter d_{model}

(2) 编码过程中的掩码长度 k 。如图 8 所示，将掩码长度 k 设置为路径的 50% 时，TRMS 在 6 个子轨迹数据集上都取得了最佳性能。实际上，设置 k 为 50% 可以平衡 Encoder 和 Decoder 的性能，这是因为预训练任务需要 Encoder 根据未掩码片段提取轨迹语义特征，然后训练 Decoder 预测掩码片段。如果 Encoder 一侧或 Decoder 一侧的有效路段太少，则会导致 MSN 将更多的注意力集中在某一侧，这对 TRL 任务不利。

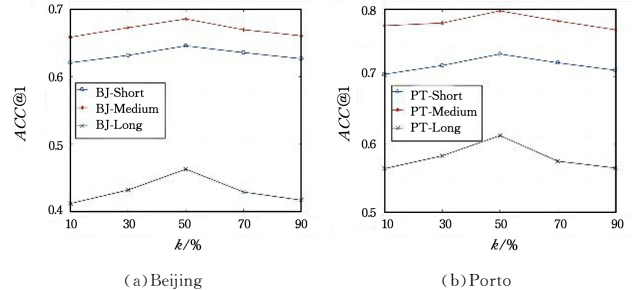


图 8 参数 k 的分析

Fig. 8 Analysis of parameter k

结束语 本文提出了一种新模型 TRMS，用于有效表征路网场景下不同长度的轨迹，该模型由以下两个模块组成：(1) TDW 通过基于概率感知游走来优化传统 DeepWalk 算法，以充分挖掘路网的拓扑结构；(2) MSN 结合 Masked Seq2Seq 学习框架和自注意力机制的优势，有效刻画轨迹的长期语义特征。最后，在真实轨迹数据集上进行实验，结果表明，TRMS 在短、中、长轨迹建模方面都优于现有最优的基线方法。在接下来的工作中，将试图引入外部数据信息，进一步探索路网中相邻路段间潜在的内部联系，以进行更精确有效的轨迹表示学习的相关研究。

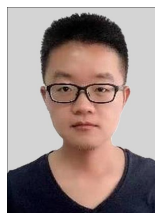
参考文献

- [1] WANG S, BAO Z, XIE Z, et al. Torch: A search engine for trajectory data[C]// Proceedings of SIGIR. ACM, 2018: 535-544.
- [2] YOU D. Trajectory Pattern Construction and Next Location Prediction of Individual Human Mobility with Deep Learning Models [J]. Computing in Science and Engineering, 2020, 14(2): 52-65.
- [3] ZHAO J, XU J. On prediction of user destination by sub-trajectory understanding: A deep learning based approach[C]// Proceedings of CIKM. ACM, 2018: 1413-1422.
- [4] LI X. Deep representation learning for trajectory similarity com-

- putation[C]//Proceedings of ICDE. IEEE,2018;617-628.
- [5] ENCISO-RODAS L. Trajectory anomaly detection based on similarity analysis[C]//Proceedings of CLEI. IEEE,2021;1-10.
- [6] WANG S,BAO Z,CONG G, et al. A survey on trajectory data management,analytics,and learning [J]. ACM Computing Surveys,2021,54(2):1-36.
- [7] ZHENG Y,LI Q,CHEN Y,et al. Understanding mobility based on GPS data[C]// Proceedings of UbiComp. ACM,2008;312-321.
- [8] ZHENG Y,LIU L,WANG L,et al. Learning transportation mode from raw GPS data for geographic applications on the web [C]//Proceedings of WWW. ACM,2008;247-256.
- [9] WU H,CHEN Z,SUN W,et al. Modeling trajectories with recurrent neural networks [C] // Proceedings of IJCAI. IEEE,2017;3083-3090.
- [10] FU T,LEE W. Trembr:Exploring road networks for trajectory representation learning [J]. ACM Transactions on Intelligent Systems and Technology,2020,11(1):1-25.
- [11] CHEN Y,CONG G,LIU Y, et al. Robust road network representation learning:When traffic patterns meet traveling semantics[C]//Proceedings of CIKM. ACM,2021;211-220.
- [12] WU N,ZHAO W,WANG J, et al. Learning effective road network representation with hierarchical graph neural networks [C]//Proceedings of KDD. ACM,2020;6-14.
- [13] JEPSEN T,JENSEN C. Graph convolutional networks for road networks[C]// Proceedings of SIGSPATIAL. ACM,2019;460-463.
- [14] LEE W,DU Y. Learning embeddings of intersections on road networks[C]// Proceedings of SIGSPATIAL. ACM,2019;309-318.
- [15] PEROZZI B,SKIENA S. Deepwalk:online learning of social representations[C] // Proceedings of KDD. ACM,2014;701-710.
- [16] SONG K,TAN X,QIN T,et al. MASS:masked sequence to sequence pre-training for language generation[C]//Proceedings of ICML. JMLR,2019;5926-5936.
- [17] GROVER A. node2vec: Scalable feature learning for networks [C]//Proceedings of SIGKDD. ACM,2016;855-864.
- [18] MIKOLOV T. Distributed representations of sentences and documents[C]//Proceedings of ICML. JMLR,2014;1188-1196.
- [19] CHANG M,LEE K. BERT:pre-training of deep bidirectional transformers for language understanding [C] // Proceedings of NAACL-HLT. ACL,2019;4171-4186.
- [20] VASWANI A,SHAZEER N,JONES L. Attention is all you need[C]//Proceedings of NIPS. ACM,2017;5998-6008.
- [21] RANU S,TELANG A. Indexing and matching trajectories under inconsistent sampling rates [C] // Proceedings of ICDE. IEEE,2015;999-1010.
- [22] ZHANG C. Map-matching for low-sampling-rate GPS trajectories[C]//Proceedings of SIGSPATIAL. ACM,2009;352-361.
- [23] YANG C. Enhanced Map-Matching Algorithm with a Hidden Markov Model for Mobile Phone Positioning [J]. International Journal of Geographical Information Science,2017,6(11):327-343.
- [24] VULIC I. Hello,it's GPT-2-how can I help you? towards the use of pretrained language models for task-oriented dialogue systems[C]//Proceedings of IJCNLP. ACL,2019;15-22.



CHEN Jiajun, born in 1998, postgraduate, is a member of China Computer Federation. His main research interests include trajectory representation learning and deep learning.



CHEN Wei, born in 1989, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include data mining, recommendation systems and knowledge graph.

(责任编辑:喻黎)